Aligning molecular data to standardized clinical terminologies will support biologically meaningful analysis of medical record data, which can be achieved by integrating external sources of biomedical knowledge.

**MAPPING CHALLENGES**
- Limitations of existing work in this domain:
  - Focused on specific diseases and biological domains
  - Largely limited to one-to-one mappings
  - Rarely include external validation
- Existing algorithms cannot automatically capture complex biological semantics underlying clinical concepts

**GOAL:** Develop OMOP2OBO, the first health system-wide integration and alignment between Observational Medical Outcomes Partnership (OMOP) standardized clinical terminologies and OBO ontologies.

**VALIDATION**
- OMOP-normalized Children’s Hospital Colorado EHR data
- OBOs were selected by domain experts and included diseases, phenotypes, anatomical entities, cell types, organisms, small molecules, vaccines, and proteins
- Mappings were performed using the pipeline in Figure 2.
- 20% of the most challenging mappings were verified by a panel of clinical and molecular domain experts.
- Mapping generalizability was assessed by comparing the coverage of mapped concepts to 24 independent EHRs.

**PRESENTATION: Tiffany J. Callahan**

**AUTHORS:**
- Tiffany J. Callahan, MPH
- Jordan M. Wyrwa, DO
- Nicole A Vasilevsky, PhD
- Peter N. Robinson, MD, PhD
- Melissa A Haendel, PhD
- Lawrence E. Hunter, PhD
- Michael G. Kahn, MD, PhD

**OMOP2OBO** is the first health system-wide resource to provide interoperability between 105020 OMOP clinical concepts and 142249 concepts in eight OBO ontologies.

**FUTURE WORK:** We are currently working on expanding the mapping provenance to include mechanisms of actions and integrating with basic science data and clinical research (Figure 1).

**Figure 1.** A Knowledge representation demonstrating how different OMOP clinical domains (i.e. conditions, drug ingredients, measurements, and immunizations) can be linked to biological mechanisms of human disease using biomedical ontologies.

**Figure 2.** An overview of the OMOP2OBO mapping algorithm. There are two primary mapping strategies: Automatic and manual. The automatic approach uses all OMOP standard concepts, ancestors, labels, and synonyms and all ontology labels, synonyms, definitions, and database cross-references.

**Figure 3.** Mapped concepts for each ontology by clinical domain (i.e. conditions, drug ingredients, and measurements) and mapping category. HPO (Human Phenotype Ontology), MONDO (Mondo Disease Ontology), OHP2OBO (Chemical Entities of Biological Interest), PR (Protein Ontology), NCBITaxon (NCBI Organism Taxonomy), VGO (Vaccine Ontology), UBERON (Uber-Anatomy Ontology), CL (Cell Ontology).